

# بررسی کیفیت داده‌های مکانی مردم‌گستر در مدیریت حوادث از منظر برازندگی استفاده، مطالعه موردی: مسیریابی

زهرا پاشائی\*<sup>۱</sup>، محمدرضا ملک<sup>۲</sup>

<sup>۱</sup> دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه

نصیرالدین طوسی

z\_pashaei69@yahoo.com

<sup>۲</sup> دانشیار دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی

mrmalek@kntu.ac.ir

(تاریخ دریافت مهر ۱۳۹۶، تاریخ تصویب اسفند ۱۳۹۶)

## چکیده

داده‌های مکانی مناسب یکی از مهم‌ترین نیازهای اساسی مدیران و برنامه‌ریزان در حوزه مدیریت حوادث است. در طی چندسال گذشته، داده‌های مکانی مردم‌گستر<sup>۱</sup> بدلیل فراهم آوردن اطلاعات بهنگام، در دسترس و ارزان، جایگاه به‌خصوصی پیدا کرد به طوری که گاهی تنها منبع اطلاعاتی در دسترس به حساب می‌آید. اما عدم آگاهی از کیفیت این دسته از اطلاعات، منجر به تردید در بهره‌وری از آن‌ها گردید. به همین دلیل بررسی کیفیت این داده‌ها به لحاظ مناسب بودن این اطلاعات در حوزه مدیریت حوادث ضروری است. از طرفی، مسیریابی گروه‌های امداد و نجات یکی از نیازهای مهم در این حوزه به‌شمار می‌آید و به عنوان مطالعه موردی انتخاب گردید و برازندگی استفاده<sup>۲</sup> به عنوان وجه قالب کیفیت که به نیاز کاربران و هدف در نظر گرفته شده بستگی دارد، مورد بررسی قرار گرفت. بنابراین در مقاله حاضر، بعنوان نمونه موردی، برآورد برازندگی استفاده داده‌های مکانی مردم‌گستر در مسیریابی انتخاب شده است. بدین ترتیب ابتدا اطلاعات هندسی و توصیفی مورد نیاز شناسایی و سپس معیارهای کامل بودن، دقت هندسی، سازگاری منطقی، دقت توصیفی و کامل بودن توصیفی نام عارضه انتخاب و برآورد شدند. در مرحله بعد به منظور برآورد کمی میزان برازندگی استفاده، استفاده از سیستم خبره فازی برای تلفیق سایر معیارها پیشنهاد گردید. در این پژوهش، منطقه ۱۱ شهر تهران به عنوان ناحیه مطالعاتی انتخاب شد و از داده‌های سایت OSM به عنوان مجموعه داده مردم‌گستر استفاده گردید. همینطور داده‌های تولیدشده توسط شهرداری به عنوان مجموعه داده‌های استاندارد بکار گرفته شد. نتایج حاصل از این مقاله، نشان داد ۷۴/۴ درصد از کل منطقه دارای برازندگی استفاده بین (۰/۹۳- ۰/۹۱) هستند و مابقی در بازه (۰/۵ - ۰/۹۱) تغییر می‌کنند و متوسط برازندگی برای کل منطقه، ۸۹ درصد به دست آمد. در ادامه، آنالیز کوتاه‌ترین مسیر به عنوان معیاری برای ارزیابی، در دو مجموعه داده اعمال شد. نتیجه شباهت ۹۵ درصدی داده‌های OSM را در قیاس با داده‌های استاندارد نشان می‌دهد.

**واژگان کلیدی:** اطلاعات مکانی مردم‌گستر، مدیریت حوادث، مسیریابی، معیارهای کیفیت، برازندگی استفاده

\* نویسنده رابط

<sup>۱</sup> Volunteered Geographic Information

<sup>۲</sup> Fitness For Use

## ۱- مقدمه

حادثه یک رویداد خارجی مشخص، قابل شناسایی، ناگه و غیرقابل پیش‌بینی، غیرعادی و بدون قصد است که در یک زمان و مکان ویژه رخ می‌دهد که منجر به خسارات زیست محیطی، مادی و بشری می‌شود. سوانح و حوادث در هر جامعه‌ای اتفاق می‌افتد که آمادگی برای مقابله با آن از وظایف مدیران و برنامه‌ریزان حوادث است. مهم‌ترین نیاز مدیران و برنامه‌ریزان در این حوزه، اطلاع دقیق از شرایط موجود است، بنابراین یکی از مهم‌ترین داده‌های موردنیاز، داده‌های مکانی صحیح و به موقع است. به عنوان مثال اطلاع از تعداد ساختمان‌های تخریب‌شده بدون مشخص بودن مکان آن‌ها از ارزش کمی برخوردار است [۱].

تولید داده‌های مکانی، برای مدت‌زمان طولانی، بر عهده سازمان‌های نقشه‌برداری ملی<sup>۱</sup> و یا شرکت‌های تجاری بزرگ فعال در عرصه جغرافیایی بوده است که شامل فرآیند پیچیده "جمع‌آوری، مدل‌سازی، مدیریت و به‌روزرسانی مجموعه داده‌های مکانی است" [۲]. در این فرآیند، به‌منظور دستیابی به بهترین کیفیت ممکن، روال استانداردسازی با دقت تمام دنبال شده و در نتیجه داده‌های تولیدشده دارای سطح بالایی از اطمینان هستند. باوجوداین، به دلیل هزینه‌بر بودن فرآیند جمع‌آوری و صدور مجوز، دسترسی به این‌گونه از اطلاعات معمولاً سخت و دشوار است.

در طی چند سال گذشته، پیشرفت فناوری‌های GPS، web2، تعبیه‌سازنده‌های تعیین موقعیت درگوشی‌های تلفن همراه و بسیاری دیگر، زمینه ایجاد گونه جدیدی از اطلاعات مکانی را فراهم آورده است، به این دسته جدید از اطلاعات، اطلاعات مکانی مردم‌گستر (VGI) گفته می‌شود که توسط عموم مردم و بدون نیاز به مهارت تخصصی، تولید و از طریق شبکه‌های اینترنتی به اشتراک گذاشته می‌شود. به‌طورکلی، دانش محلی، بهنگام بودن، دسترسی رایگان به داده‌ها، برخی از خصوصیات بارز داده‌های VGI است که آن‌ها را از مجموعه داده‌های معتبر متمایز می‌سازد [۳].

در میان فعالیت‌های VGI مبنای Open Street Map (OSM)، یکی از محبوب‌ترین و مشهورترین فعالیت‌هایی

است که هدف آن ارائه نقشه قابل‌ویرایش از کل جهان است. توسعه موفقیت‌آمیز محتوای تولیدشده توسط کاربران در سال‌های اخیر تأثیر زیادی بر روی زمینه‌های مختلف همانند خدمات مبتنی بر مکان<sup>۲</sup>، مدل‌های سه‌بعدی و غیره گذاشته است. اما از آنجایی که کاربران این پروژه‌ها، اغلب افراد غیرمتخصص هستند و داده‌ها را بدون رعایت استانداردهای موجود جمع‌آوری می‌کنند، این امر موجب گردید، هیچ‌گونه نظارتی بر کنترل کیفیت داده‌ها وجود نداشته باشد [۴]. بنابراین عدم آگاهی از کیفیت داده‌های VGI، یکی از محدودیت‌های این داده‌ها تلقی می‌گردد و ضروری است که بدانیم آیا VGI به‌عنوان یک جایگزین ارزان و کم‌هزینه، علی‌رغم عدم آگاهی از کیفیت آن، می‌تواند با داده‌های مکانی دارای قابلیت اطمینان بالا، رقابت کند. پاسخ به این سؤال مستلزم این است که پیش از استفاده از داده‌های VGI، نسبت به کیفیت آن اطمینان حاصل کنیم.

معمولاً ارزیابی کیفیت از طریق برآورد شاخص‌های آن امکان‌پذیر است، در میان معیارهای کیفیت، شاخص برازندگی استفاده به دلیل تأثیرپذیری از سایر شاخص‌های کیفیت، به‌عنوان مهم‌ترین معیار و وجه قالب کیفیت تلقی می‌گردد و گاهی با عنوان کیفیت بیرونی و یا عملی از این شاخص یاد می‌شود [۵]. در واقع شاخص برازندگی استفاده، میزان تناسب یک مجموعه داده، برای کاربردی خاص را بیان می‌کند. به‌طور مثال ممکن است یک پایگاه داده برای کاربردی خاص مناسب و برای کاربردی دیگر مناسب نباشد. بنابراین ارزش یک مجموعه داده از کاربردی به کاربرد دیگر و از هدفی به هدف دیگر متفاوت است [۶].

از طرف دیگر، در هنگام رخداد حوادثی نظیر حادثه آتش‌سوزی و ریزش ساختمان پلاسکو و برج گرانفل<sup>۳</sup>، مسئله مسیریابی گروه‌های امداد و نجات برای انتقال مجروحین به نقاط امن، یکی از نیازهای اساسی محسوب می‌شود. از این رو، مسیریابی به عنوان نمونه موردی موردنظر ما در حوزه مدیریت حوادث انتخاب شد. به عبارتی دیگر، هدف از پژوهش حاضر پاسخ به این سوال است که آیا داده‌های مکانی مردم‌گستر توانایی برآورده کردن نیازهای مکانی در کاربرد خاص مسیریابی را دارند.

<sup>۲</sup> Location Based Services

<sup>۳</sup> Grenfell Tower

<sup>۱</sup> National Mapping Agencies

بودن، دقت هندسی و دقت توصیفی را ارزیابی کرد. نتایج این تحقیق، کامل بودن ۸۱٪، متوسط دقت هندسی ۳/۸ متر و دقت توصیفی ۶۷٪ را نشان داد. Brovelli و همکاران [۱۰] کیفیت ساختمان‌های OSM را در منطقه-ای به وسعت ۱۸۰ کیلومترمربع در کشور ایتالیا مورد بررسی قرار دادند. در این تحقیق برخلاف بسیاری از تحقیقات که بر روی کامل بودن ساختمان‌های OSM تمرکز کرده‌اند، دقت هندسی نیز مورد ارزیابی قرار گرفت. در این تحقیق برای نمایش عدم تجانس عوارض OSM، از یک روش پیکسل مبنا استفاده شده است. همچنین دقت هندسی با استفاده از فن شناسایی زوج‌های همسان<sup>۱</sup> برآورد گردید. نتایج این تحقیق، بهبود دقت هندسی و کاهش نسبی کاملی از مرکز به پیرامون شهر را نشان داد. از جمله تحقیقات صورت گرفته در کشور ایران، می‌توان به تحقیق صورت گرفته توسط محمدی و ملک [۱۱، ۱۲] اشاره کرد که در ابتدا با استفاده از یک روش جدید نقطه مبنا به تناظر یابی دو مجموعه داده OSM و استاندارد پرداختند و سپس به منظور بهبود سازگاری منطقی داده-های OSM، از یک روش ترکیبی بر مبنای هوش مصنوعی برای محاسبه دقت هندسی عوارض بدون متناظر استفاده کردند. نمونه‌های دیگر این دسته از تحقیقات را می‌توان در کار [۱۳-۱۵] نیز مشاهده کرد.

دسته‌ای دیگر از تحقیقات، برخلاف روش اول، داده-های مردم‌گستر را با استفاده از ویژگی‌های ذاتی خود داده‌ها و بدون مقایسه موردسنجش قرار می‌دهند که خود به چندین روش داده مبنا، کاربر مبنا و رابطه مبنا تقسیم شده است. به طور مثال، Ciepluch و همکاران [۱۶] شاخص‌هایی نظیر توزیع مکانی عوارض، تعداد عوارض و نقاط را برای برآورد کیفیت داده‌های VGI مبنا پیشنهاد کردند. حسینی و عباسپور [۱۷] با استفاده از درخت تصمیم و شبکه عصبی، به تحلیل میزان دقت دسته‌بندی خیابان‌ها، به عنوان بخشی از دقت معنایی پرداختند.

در میان تحقیقات صورت گرفته اغلب معیارهای کیفیت نظیر دقت مکانی، کامل بودن و دقت موضوعی ارزیابی شده‌اند اما وجه قالب کیفیت یعنی برازندگی استفاده به صورت کلی بحث شده است. بنا به تحقیقات صورت گرفته در زمینه این شاخص، می‌توان آن را به سه

ساختار مطالبی که در ادامه خواهد آمد به این شرح است که در بخش دوم به مرور تعدادی از تحقیقات مرتبط در این زمینه می‌پردازیم، در بخش سوم مبانی نظری پژوهش حاضر توضیح داده می‌شود، بخش چهارم رویکرد پیشنهادی و روش‌های مورداستفاده بحث خواهد شد، بخش پنجم به پیاده‌سازی روش پیشنهادی و ارزیابی پرداخته و در بخش آخر نتیجه‌گیری ارائه خواهد شد.

## ۲- پیشینه تحقیق

موضوع ارزیابی کیفیت داده‌های مکانی مردم‌گستر چندین سال است که توجه بسیاری از محققان را به خود معطوف ساخته است و در این راستا تحقیقات بسیار زیادی انجام شده است. یکی از راهکارهای رایج، مقایسه داده‌های مکانی مردم‌گستر با مجموعه داده استاندارد و یا مرجع است. در این دسته از تحقیقات اغلب بعد از اتمام فرآیند تناظر یابی، معیارهای تعریف شده توسط استاندارد ISO بررسی می‌گردند. به طور مثال Fan و همکاران [۷] در یک فرآیند سه مرحله‌ای شامل: یافتن داده‌های نظیر، محاسبه شاخص‌های کیفیت و آنالیز آماری، کیفیت عوارض مسطحاتی OSM را در مقایسه با داده‌های استاندارد ATKIS محاسبه کردند. آن‌ها شاخص‌های دقت معنایی، دقت مکانی، کامل بودن و صحت شکل را مورد ارزیابی قرار دادند و نتایج کامل بودن بالای داده‌های OSM، دقت معنایی ۱۰۰٪، شباهت بالای دو مجموعه و وجود انحراف معیار ۴ متری در دقت هندسی را نشان داد. Seibriz [۸] کیفیت داده‌های OSM را در مقایسه با مجموعه مرجع برای ایالات مختلف آفریقا مورد بررسی قرار داد. هدف از برآورد کیفیت در این تحقیق، بررسی توانایی داده‌های OSM در تلفیق با داده‌های مرجع است. همچنین با فرض پذیرفتن سازگاری OSM با داده‌های مرجع و توانایی آن در به‌روزرسانی نقشه‌های موجود، معیارهای دقت هندسی، دقت معنایی، کامل بودن عوارض خطی و دقت موضوعی عوارض مسطحاتی بررسی و سپس معیارهای کیفی نظیر به روز بودن عوارض خطی و مسطحاتی ارزیابی شدند. نتایج به دست آمده حاکی از رد فرضیه این تحقیق است. Will [۹] کیفیت عوارض خطی OSM در سوئد را بررسی کرد. او در ابتدا روش عارضه‌مبنای مطرح شده توسط Kouklotsous [۲۷] را بهبود و سپس معیارهای کامل

<sup>۱</sup> Homologous Pairs Detection

و اما در زمینه مدیریت حوادث تحقیقاتی به بررسی اعتبار داده‌های مکانی مردم‌گستر پرداختند که می‌توان به تحقیق صورت گرفته توسط Poser و همکاران [۲۱] اشاره کرد. در این تحقیق، میزان اعتبار داده‌های گردآوری شده توسط مردم را با داده‌های اندازه‌گیری شده برای سیل سال ۲۰۰۲ در ایالت متحده الدنبرگ<sup>۲</sup> مقایسه کردند. در این پژوهش، این دو مجموعه داده به عنوان ورودی، وارد مدل تخمین خسارت سیل<sup>۳</sup> شده تا میزان خسارت وارد بر ساختمان‌های مسکونی برآورد گردد. سپس نتایج تخمین خسارات این مدل با نتایج بدست آمده از مدلسازی هیدرولیکی مقایسه شدند. نتایج این تحقیق نشان داد که مدلسازی خسارات با استفاده از داده‌هایی که از طریق مردم گردآوری شده به نتایج مدلسازی هیدرولیکی نزدیک‌تر است.

Hung و همکاران [۲۲]، اعتبار داده‌های مردم‌گستر را در هنگام پاسخگویی به بحران سیل در کشور استرالیا بررسی کرده‌اند. آن‌ها مجموعه داده‌های دو رویداد سیل بریسیین<sup>۴</sup> در سال‌های ۲۰۱۱ و ۲۰۱۳ را جمع‌آوری و مطابق با فاکتورهای مکانی و استفاده از داده‌های سیل سال ۲۰۱۱ یک رگرسیون منطقی باینری برای تعیین میزان اعتبار داده‌های مردم‌گستر ساختند. سپس عملکرد این مدل احتمالی با استفاده از نمونه‌های تست سال ۲۰۱۳ مورد ارزیابی قرار گرفت. آنها به این نتیجه رسیدند که مدل، پتانسیل بالایی دارد و می‌تواند اعتبار داده‌های مردم‌گستر را برای پاسخ موثر و سریع تصمیم‌گیری و هماهنگی در مدیریت بحران دسته‌بندی کند.

همانطور که ملاحظه گردید، ابتدا تحقیقاتی که در زمینه کیفیت داده‌های VGI انجام شده مرور و سپس نمونه‌ای از تحقیقاتی که در زمینه مدیریت بحران و داده‌های مردم‌گستر انجام شده، اشاره شد. تحقیقاتی که در زمینه مدیریت بحران و VGI صورت گرفته، اغلب به بررسی اعتبار این داده پرداخته‌اند. اما دسته اول تحقیقات که در زمینه برآورد کیفیت داده‌های مکانی مردم‌گستر صورت گرفته است، تنها به بررسی معیارهای درونی کیفیت نظیر دقت مکانی، کامل بودن، دقت معنایی اکتفا کرده‌اند.

دسته بررسی برازندگی استفاده از جنبه بهره‌وری بیشینه از داده، برآورد کمی و بررسی مناسب بودن داده‌های VGI در کاربرد خاص در نظر گرفت.

بهره‌وری بیشینه از داده مبتنی بر این است که از روش‌های مناسب برای ذخیره‌سازی و پردازش بهینه داده‌ها استفاده گردد. به‌کارگیری برچسب مناسب راهکاری است که تعدادی از محققان پیشنهاد کرده‌اند. به‌طور مثال ایجاد یک سیستم توصیه‌گر OSMANTIC توسط Vandecasteele و Devillers [۱۸] به کاربران این امکان را فراهم می‌کند تا برچسب مناسبی در هنگام ورود داده انتخاب نمایند. بعلاوه عشقی و آل شیخ [۱۹] یک مدل مفهومی الهام گرفته‌شده از مدل CityGML را ایجاد کردند. بدین ترتیب هنگام ذخیره‌سازی داده در پایگاه داده، علاوه بر کاربرد موردنظر کاربر، سایر کاربردهای دیگر نیز در نظر گرفته‌شده است و داده ورودی به‌صورت بهینه‌تر ذخیره می‌گردد.

دسته‌ای دیگر، با بهره‌گیری از مفاهیم کیفیت درونی، بیرونی، به برآورد کیفیت عملی پرداختند. به‌طور مثال، Bordogna و همکاران [۵]، واحدی [۶] با استفاده از کمیت سنج‌های زبانی و روش‌های تصمیم‌گیری چندمعیاره OWA<sup>۱</sup>، کیفیت عملی داده‌های مکانی مردم‌گستر را به ترتیب در مطالعه یخچال‌های طبیعی و بناهای تاریخی ارزیابی کردند.

و اما دسته آخر، علی‌رغم جستجوی گسترده نگارنگان، تنها یک تحقیق در این زمینه یافت شد. این تحقیق توسط Mondzech و Sester [۲۰] انجام شده به طوری که برای یافتن مناسب‌ترین پایگاه داده در ناوبری پیاده‌رو، دو پایگاه داده OSM و ATKIS مقایسه شدند. در این تحقیق از الگوریتم کوتاه‌ترین مسیر جهت مقایسه دو مجموعه داده استفاده و نتیجه حاصل این‌گونه توصیف شد؛ پایگاه داده‌ای مناسب است که طول حاصل از اتصال دو نقطه ابتدا و انتهای آن کوتاهتر باشد. با وجود اینکه مناسب بودن داده‌های مردم‌گستر در این تحقیق بررسی شده اما نتوانسته برآورد کمی از این معیار ارائه دهد. درحالی‌که در تحقیق حاضر این معیار به صورت کمی ارائه شده است.

<sup>۲</sup> Eildenburg

<sup>۳</sup> Flood Loss Estimation

<sup>۴</sup> Brisbane

<sup>۱</sup> Ordered Weighted Average

به گروه‌های مختلفی طبقه‌بندی می‌شوند. در ادامه به تعریف کوتاهی از هر یک از معیارهای کیفیت می‌پردازیم. کامل بودن؛ معیاری برای سنجش حضور و یا عدم حضور، ویژگی‌ها و روابط موجود بین عوارض است. سازگاری منطقی؛ معیاری است برای سنجش میزان وابستگی میان قوانین منطقی ساختار داده، ویژگی و روابط میان عوارض. دقت هندسی؛ به‌عنوان معیاری برای بررسی دقت موقعیت عوارض در یک سیستم استاندارد مکانی تعریف می‌شود. دقت موضوعی؛ این شاخص معیاری است برای سنجش دقت توصیفات کمی، درستی توصیفات غیر کمی و درستی طبقه‌بندی و روابط عوارض. دقت زمانی؛ این شاخص، توصیفات و روابط زمانی عوارض را تعریف می‌کند.

کیفیت بیرونی، میزان تطابق بین یک محصول و نیازهای کاربران را بیان می‌کند به‌عبارتی دیگر کیفیت یک محصول یک مقدار مطلق نیست و یک محصول یکسان ممکن است کیفیت متفاوتی برای کاربران مختلف داشته باشد. ISO [۲۴] کیفیت بیرونی را این‌گونه تعریف می‌کند: "تمام ویژگی‌ها و مشخصات یک محصول که توانایی لازم برای برآورده کردن نیازهای خواسته‌شده را داشته باشد." کیفیت بیرونی اغلب به‌عنوان برآوردی استفاده و یا برآوردی برای هدف تعریف می‌شود.

جالب این است بدانیم که کیفیت بیرونی می‌تواند از طریق معیارهایی که کیفیت درونی داده را توصیف می‌کند نیز ارزیابی گردد. به‌طور مثال، برای اینکه ببینیم آیا یک مجموعه داده با نیازهای کاربر مطابقت دارد، نه‌تنها آن مجموعه داده از نظر عوارض و ویژگی‌های ضروری می‌بایست بررسی گردد، بلکه باید دید که آیا به لحاظ هندسی دقت کافی را دارد. درواقع با این کار معیارهای درونی کیفیت توصیف می‌شوند. اما نکته حائز اهمیت این است که ارزیابی کیفیت بیرونی داده برخلاف کیفیت درونی داده هنوز هم یک زمینه بسیار ناب برای تحقیق و تفسیر باز گذاشته است.

### ۳-۲- سیستم خبره فازی

منطق فازی نظریه ریاضی مرسوم برای بیان پیچیدگی، عدم حتمیت و مسائل غیر ساختاری است. در مقایسه با منطق کلاسیک که تنها به‌درستی یا نادرستی ارزش می‌پردازد، منطق فازی درجه‌های متفاوتی از درستی را می‌تواند بیان کند [۲۵]. همان‌طور که در شکل ۲ نشان

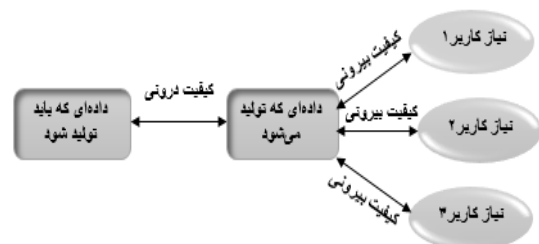
درحالی‌که برآوردی استفاده به‌عنوان کیفیت بیرونی به صورت کلی و فقط در غالب پیشنهادات ارائه شده است. از طرفی تحقیقات در زمینه برآوردی استفاده، بیشتر از جنبه بهبود ذخیره‌سازی داده‌هایی غیر از داده‌های موجود انجام شده است. درحالی‌که برآوردی استفاده‌های موجود از نظر مناسب بودن یک مجموعه داده جهت برآورده کردن نیازهای یک کاربرد مشخص کمتر مورد توجه قرار گرفته‌اند. بنابراین نوع‌آوری تحقیق حاضر ارائه رویکردی برای بررسی برآوردی داده‌های مردم‌گستر در حوزه مدیریت حوادث و به طور خاص مسیریابی گروه‌های امدادسانی است.

### ۳-۱- مبانی نظری

در این بخش، مفاهیم مربوط به کیفیت داده‌های مکانی مردم‌گستر و سیستم خبره توضیح داده شده است.

#### ۳-۱- کیفیت داده‌های مکانی مردم‌گستر

دیدگاه‌های متفاوتی از مفهوم کیفیت داده‌های مکانی وجود دارد، عده‌ای متصورند داده‌ای که در عمل تولید می‌شود، اگر به لحاظ استاندارد کامل باشد و یا به عبارتی هیچ خطایی نداشته باشد داده‌ای باکیفیت بالا است درحالی‌که عده‌ای دیگر بر این باورند که اگر داده‌ای که تولید می‌شود با نیازهای سازمان متقاضی داده یا کاربر موردنظر مطابقت داشته باشد داده‌ای باکیفیت بالا خواهد بود [۱۲]. همان‌طور که از شکل ۱ پیداست بسیاری از محققان کیفیت داده‌ای مکانی را به دو گروه کیفیت درونی و کیفیت بیرونی تقسیم کرده‌اند.



شکل ۱- دسته‌بندی کیفیت درونی و بیرونی [۲۳]

کیفیت درونی سطح تشابه میان داده تولیدشده و داده‌ای که می‌بایست تولید شود را بیان می‌کند. کیفیت درونی داده با استفاده از شاخص و یا معیارهای کیفیت توصیف می‌گردد [۲۳]. درواقع شاخص‌های کیفیت، یک مؤلفه برای توصیف جنبه مشخصی از کیفیت داده‌های مکانی است که

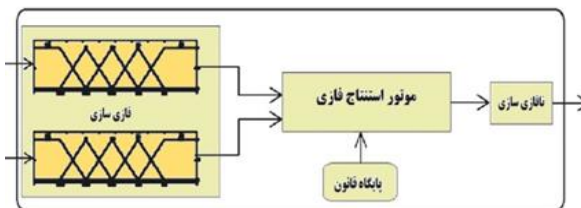
است. ارقام توصیفی نیز شامل نام عارضه، طول، عرض، جهت و نوع عارضه هستند که در تحقیق حاضر، نام عارضه به‌عنوان رایج‌ترین قلم توصیفی شبکه راه به دلایلی که ذکر خواهد شد انتخاب می‌گردد. اولاً؛ نوع عارضه راه به دلیل تفاوت در طبقه‌بندی استفاده‌شده در هر مجموعه داده، معمولاً ارزیابی آن با مشکل روبروست. دوما؛ یافتن سریع آدرس محل حادثه و یا انتقال مجروحین به بیمارستان‌ها و موارد مشابه با این موضوع یکی از اهداف اصلی راه‌یابی و مسیریابی گروه‌های امدادونجات و آتشنشانان تلقی می‌گردد. به طور مثال جهت یافتن "اتوبان همت"، اگر روی نقشه "اتوبان همت" یا "بزرگراه همت" جستجو شود شاید تفاوتی در جستجو بوجود نیآورد. درحالی‌که به طور مثال "اتوبان حکیم" با "اتوبان همت" متفاوت است. هرچند در برخی موارد این مورد نقض می‌گردد. اما بعد از نظرخواهی از کارشناسان در این مورد خاص، به این نتیجه رسیدیم که آگاهی از نام عارضه راه در قیاس با نوع عارضه از اهمیت بیشتری برخوردار است و می‌تواند در پیدا کردن مسیر صحیح یکی از اطلاعات مهم به حساب آید. بنابراین به بررسی کیفیت اطلاعات توصیفی از نوع مقیاس اسمی پرداخته شد.

#### ۴-۲- تعیین معیارهای کیفیت

شاخص برازندگی استفاده اساساً به نیاز کاربران، کاربرد و هدف در نظر گرفته‌شده برای پروژه وابستگی شدیدی دارد، از این رو استخراج معیارهای تأثیرگذار بر کیفیت داده‌های مکانی مردم‌گستر در کاربرد خاص، از مهم‌ترین بخش این پژوهش به حساب می‌آید. اساس انتخاب معیارهای تعریف‌شده، معیارهای پیشنهادشده در تحقیقات پیشین است. به‌طور مثال بسیاری از محققان کامل بودن و دقیق بودن شبکه راه را معیارهای مهم و کلیدی در این حوزه شناسایی کرده‌اند [۲۷].

همچنین Grasser و همکاران [۲۸] معیارهای یک‌طرفه بودن و محدودیت دور زدن ۱ را به‌عنوان معیارهای جدید در این زمینه معرفی کردند. اما در این میان، جامع‌ترین و کامل‌ترین معیارهای تعریف‌شده در کاربردهای مرتبط با سرویس‌های مکان‌مبنا را می‌توان در تحقیق صورت گرفته توسط Barran و همکاران [۲۹] مشاهده کرد. آن‌ها بیش از

داده‌شده است، فرآیند ارزیابی استنتاج منطق فازی شامل سه مرحله: فازی سازی، استنتاج و نافازی سازی است.



شکل ۲- قسمت‌های مختلف یک سیستم فازی [۳۷]

فازی سازی، به‌عنوان نگاشتی از یک نقطه به یک مجموعه فازی تعریف می‌شود که بوسیله تعیین توابع عضویت امکان پذیر است. در روش استنتاج فازی، مهم‌ترین بخش ساخت پایگاه دانش است که از مجموعه‌ای قوانین برپایه متغیرهای زبانی تشکیل شده است. موتور استنتاج، منطق و اصول استدلالی است که به ترکیبی از شروط موجود در پایگاه دانش اعمال شده و نتایج فازی را منجر می‌گردد [۲۶]. روش‌های استنتاج متفاوتی وجود دارد که از رایج‌ترین آن‌ها، می‌توان روش‌های استنتاج ممدانی و سوگنو را نام برد. مرحله آخر، نافازی سازی، واحدی است که به‌صورت تابع از یک مجموعه فازی به یک مقدار قطعی عمل می‌کند.

#### ۴- روش پیشنهادی و مدلسازی

ناوبری و مسیریابی یکی از مسایل مهم در حوزه مدیریت حوادث به شمار می‌آیند که نیازمند اطلاعات مکانی مناسب و تا حد زیادی کامل هستند. معمولاً داده‌های تهیه‌شده توسط سازمان نقشه‌برداری و شهرداری جوابگوی نیازهای کارشناسان این بخش است، اما نکته حائز اهمیت اینجاست که اگر داده‌های تهیه‌شده توسط مردم جایگزین این نوع از داده‌ها شوند آیا می‌توانند نیازهای داده‌ای در این حوزه را برآورده کنند، بنابراین ارائه رویکردی که بتواند میزان برازندگی اطلاعات مکانی مردم‌گستر را بررسی کند، ضروری به نظر می‌رسد. به‌طور کلی روش پیشنهادی ما مطابق شکل ۳ به چهار بخش تقسیم می‌گردد.

#### ۴-۱- تعیین داده‌های مکانی مورد نیاز

بخش اول تعیین حداقل داده‌های مکانی و توصیفی موردنیاز است. بارزترین اطلاعات هندسی موردنیاز شبکه راه است که از مجموعه‌ای از خطوط و نقاط تشکیل شده

۱ Turn Restriction

مرحله اول؛ اعمال الگوریتم Douglas-Peucker برای جنرالیزاسیون عوارض استاندارد. این مرحله یک مرحله کاملاً اختیاری است و به منظور کاهش تعداد عوارض استاندارد و یا کاهش زمان محاسبات، پیشنهاد می‌شود.

مرحله دوم؛ مجموعه عوارض موجود در مجموعه استاندارد و مردم‌گستر به بخش‌های کوچک‌تری تقسیم می‌گردند. واضح است که در این مرحله تعداد عوارض خطی افزایش خواهد یافت که برای بهبود نتایج مراحل بعدی ضروری است.

مرحله سوم؛ در این مرحله درجه عضویت هر راس از مجموعه عوارض استاندارد محاسبه می‌گردد. درجه گره هر راس از طریق شمارش تعداد یال‌های متصل به هر راس بدست می‌آید. این مرحله برای تعیین رئوس با درجه عضویت ۱ انجام می‌شود.

مرحله چهارم؛ قید فاصله جستجو با معرفی حد آستانه-ای اعمال می‌گردد. در این تحقیق مقدار فاصله تعیین شده پیش از تناظریابی، به‌عنوان حد آستانه معرفی و بافری با این مقدار اطراف مجموعه عوارض استاندارد ترسیم می‌گردد. سپس عوارض مردم‌گستر که با بافر همپوشانی دارند استخراج یافته و عوارض خارج بافر از مجموعه عوارض مردم‌گستر خارج می‌گردند. همچنین در این مرحله شکل بافر به مقدار درجه عضویت رئوس وابسته است. در صورتی که درجه عضویت ۱ باشد، بافر بدون کلاک<sup>۱</sup> ترسیم می‌شود. با این کار، طول قسمتی از عوارض مردم‌گستر که بلندتر از عارضه استاندارد تناظرش است، حذف می‌گردد.

مرحله پنجم؛ مرحله آخر راستای هر قطعه عارضه در مجموعه استاندارد محاسبه می‌گردد. مجدداً بافری اطراف مجموعه داده استاندارد ترسیم و راستا برای هر عارضه از مجموعه مردم‌گستر که در درون این بافر قرار می‌گیرند، محاسبه می‌شود. سپس راستای عوارض در دو مجموعه داده باهم مقایسه می‌شوند. اگر مقدار اختلاف به‌دست‌آمده بیش از مقدار حد آستانه معرفی شده باشد، عارضه موردنظر از مجموعه عوارض مردم‌گستر خارج می‌گردد. هدف از این مرحله، پاک‌سازی مجموعه عوارض مردم‌گستر از عوارضی است که به‌اشتباه تناظر شده‌اند و درنهایت نقشه‌ای از عوارض مردم‌گستر که دارای نظیر در مجموعه استاندارد هستند، تولید می‌شود [۳۰].

۲۵ معیار مؤثر در سنجش برازندگی استفاده در کاربردهای متفاوت، معرفی می‌کنند. در این پژوهش از معیارهای معرفی شده توسط Barran و همکاران [۲۹] به‌منظور بررسی برازندگی استفاده داده‌های مردم‌گستر استفاده شده است. مطابق با [۲۹]، معیارهای مهم در کاربرد مسیریابی و ناوبری شامل: کامل بودن شبکه راه، سازگاری منطقی، دقت هندسی شبکه راه، به‌روز بودن شبکه راه، دقت توصیفی شبکه راه و کامل بودن اسامی و یا شماره راه‌ها هستند.

#### ۴-۳- اندازه و معیارهای کیفیت استفاده شده

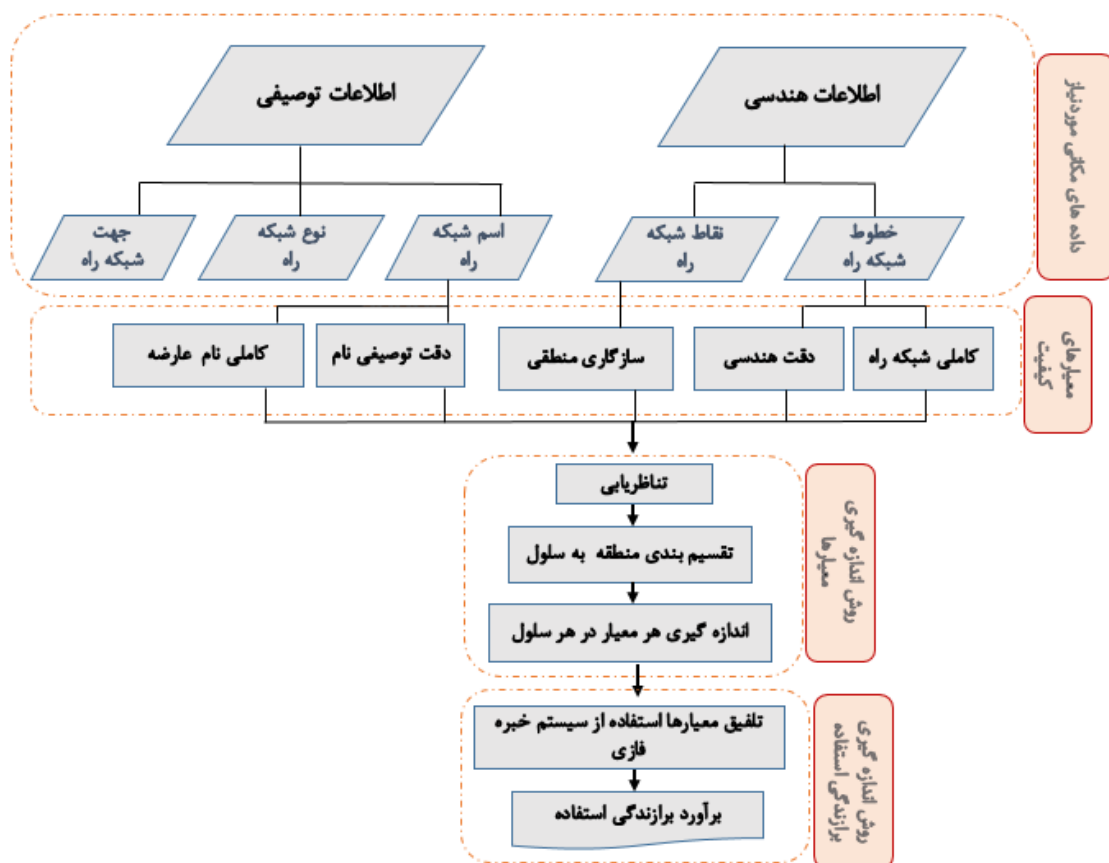
روش‌های محدودی برای برآورد معیارهای کیفیت موجود است. یکی از روش‌ها، مقایسه مجموعه داده مردم-گستر با مجموعه داده استاندارد است. در این تحقیق، برای برآورد تمام معیارها به‌جز سازگاری منطقی و کامل بودن توصیفی از روش مقایسه استفاده شده است. اولین مرحله در انجام فرآیند مقایسه، تناظر یابی بین دو مجموعه داده است به این دلیل که محاسبه میزان اختلاف، به تعیین عوارض نظیر وابسته است. در ادامه به توضیح مختصری از روش تناظر یابی اقتباس شده می‌پردازیم.

به‌طور کلی فرآیند تناظر یابی از حالت بصری و دستی به حالت خودکار توسعه یافته است. به‌طوری‌که برخی از محققان بر روی خودکارسازی فرآیند تناظر یابی تمرکز کرده‌اند و روش‌هایی نظیر روش عارضه مبنا و نقطه مبنا را توسعه دادند [۱۲، ۲۷]. از طرفی تعدادی از محققان همانند Broveli و همکاران [۳۰] فرآیند تناظر یابی را در قالب یک جعبه‌ابزار و ماژول توسعه داده‌اند. در پژوهش حاضر از ماژول توسعه داده شده توسط Brovelii برای این منظور استفاده شده است. لازم به ذکر است، بدلیل اینکه برخی مراحل فرآیند تناظریابی، نیازمند ورود مقداری جهت اعمال فاصله جستجو است. از این رو فاصله جستجوی مناسب پیش از تناظریابی تعیین می‌گردد. برای این‌منظور در ابتدا یک عرض بافر اولیه حول مجموعه عوارض استاندارد ترسیم و طول عارضه مردم‌گستر که درون این بافر قرار می‌گیرد، محاسبه می‌شود. سپس در یک فرآیندی تکراری آن‌قدر اندازه بافر تغییر می‌کند تا درصد اشتراک عارضه درون آن به‌اندازه دلخواه برسد. پس از انتخاب عرض بافر مناسب، فرآیند ارزیابی شروع می‌شود. مراحل فرآیند را در ادامه می‌آوریم.

<sup>۱</sup> Cap

توزیع یکنواخت کیفیت عوارض در بخش‌های مختلف یک منطقه است. به این مفهوم که برخی مناطق دارای کیفیت بیشتر و برخی دیگر کیفیت پایین‌تری دارند. از این رو برای نمایش بیشترین تغییرات کیفیت در مجموعه عوارض مردم‌گستر، ابتدا منطقه مورد مطالعه به سلول‌های با ابعاد یکسان تقسیم و سپس به برآورد هر یک از معیارها می‌پردازیم. لازم به ذکر است در سلول‌هایی که هیچ عارضه‌ای قرار نگرفته، از شبکه سلولی خارج می‌گردند.

بعد از اتمام فرآیند تناظریابی، نوبت به برآورد هر یک از معیارها می‌رسد. یکی از راهکارهای ارائه معیارها، بررسی آن‌ها در یک منطقه شهری است. اما از آنجایی که داده‌های مردم‌گستر ماهیت ناهمگونی دارند بنابراین بهتر است در ابعاد کوچکتری مورد ارزیابی قرار گیرند. همچنین، تقسیم‌بندی منطقه به سلول‌های کوچک در تحقیقات صورت گرفته توسط Koukoletsos [۲۷] ، Will [۹] ، Siebritz [۸] دیده می‌شود. منظور از ناهمگونی، عدم



شکل ۳- فلوجارت روش پیشنهادی

طول عوارض متناظر بر طول کل عوارض در آن مجموعه داده محاسبه گردد [۲۷].

معیار بعدی دقت هندسی است که روش‌های مختلفی برای برآورد دقت هندسی عوارض خطی وجود دارد. فاصله هاسدروف یکی از روش‌هایی است که برای اولین بار توسط Devogole و Mustiere [۳۲] پیشنهاد شد. در این روش کوتاه‌ترین فاصله از تک‌تک نقاط تشکیل‌دهنده دو خط محاسبه شده و سپس بیشترین مقادیر از مقادیر قبلی، فاصله هاسدروف را تعیین می‌کند.

کاملی شبکه راه نقش مهمی در کاربردهای مسیریابی و ناوبری بازی می‌کند [۲۹]. یک روش متداول برای برآورد کاملی شبکه راه، مقایسه طولی میان دو مجموعه استاندارد و مردم‌گستر است. در این روش مجموع طول کل عوارض در هر دو مجموعه محاسبه شده و اختلاف حاصل بیانگر کاملی شبکه راه است. اما این روش اغلب زمانی استفاده می‌گردد که بعد از انجام تناظریابی، عارضه متناظر در دو مجموعه وجود نداشته باشد [۳۱]. اما ارزیابی این معیار زمانی معنادار خواهد بود که دو مجموعه داده در ابتدا متناظر گردند و سپس از طریق نسبت جمع



در رابطه ۱،  $i$  و  $j$  به ترتیب برابر طول رشته  $a$  و  $b$  است و  $1(a_i \neq b_i)$  تابع علامت نامیده است و اگر  $a_i$  با  $b_i$  برابر باشد، تابع علامت برابر صفر و در غیر این صورت برابر ۱ خواهد بود. لازم به ذکر است که پیش از اعمال این الگوریتم، پیش پردازش‌هایی جهت استانداردسازی نام عوارض در دو مجموعه داده صورت می‌گیرد که از توضیح آن صرف نظر می‌کنیم. لازم به ذکر است، برای برآورد دقت توصیفی، تنها عوارض دارای نام مورد بررسی قرار خواهد گرفت به طوری که، فاصله ویرایش levenshtein روی تک تک عوارض اعمال و خروجی آن در فیلد مربوطه در جدول توصیفی ذخیره می‌شود. خروجی این الگوریتم مقادیری بین صفر و یا بیشینه طول بین دو رشته خواهد بود. که در پژوهش حاضر، مقادیر صفر و یک برای ما قابل قبول است. مقدار صفر به این مفهوم است که کاراکتری دو کلمه منتخب از دو مجموعه داده، کاملاً شبیه به هم هستند، درحالی که مقدار یک، بیانگر اختلاف یک کاراکتری است. مقدار یک به این دلیل قابل قبول در نظر گرفته شد که اختلاف ایجاد شده بیشتر برگرفته از فاصله ایجاد شده بین کلمات در یک مجموعه است. به طور مثال: "بوستان سعدی" در مجموعه داده مردم‌گستر و "بوستان سعدی" در مجموعه داده استاندارد.

یکی دیگر از معیارها کاملاً توصیفات است. در این مقاله، کاملاً نام عارضه مورد بررسی قرار گرفت که از طریق نسبت طول عوارضی از مجموعه داده مردم‌گستر که دارای نام هستند بر طول کل عوارض مردم‌گستر در هر سلول، قابل محاسبه است.

سازگاری منطقی شبکه راه نیز یکی دیگر از معیارهای کلیدی در کاربردهای مسیریابی است. Neis و همکاران [۳۵] سه خطای توپولوژی (شکل ۴) در نظر گرفته‌اند. ۱- خطای از هم رد شدگی و یا به هم نرسیدگی<sup>۵</sup> ۲- خطای بر روی هم افتادگی خطوط<sup>۶</sup> ۳- خطوط متقاطعی که هیچ نقطه‌ای به اشتراک نگذاشته‌اند<sup>۷</sup>. این معیار برخلاف معیارهای قبلی، بدون نیاز به مقایسه دو مجموعه داده، از طریق تعریف روابط توپولوژی قابل تعیین هستند.

Ghoochild و Hunter [۳۳] روش بافرمبنا و بافر افزایشی<sup>۱</sup> را برای برآورد این معیار معرفی می‌کنند. که گاهی برای ارزیابی کاملی شبکه راه از این روش استفاده می‌گردد. به طوری که در این روش، نسبت طول عارضه مورد تست در درون بافر عارضه مرجع متناظرش سنجیده می‌شود. از دیگر متریک و یا واحدهای اندازه‌گیری که برای این منظور به بکار می‌روند می‌توان به متوسط خطای مربعات<sup>۲</sup> ماکسیمم انحراف<sup>۳</sup> و یا فاصله اطمینان اشاره کرد. در پژوهش حاضر از ماکسیمم انحراف برای سنجش این شاخص کیفیت استفاده خواهد شد.

معیار بعدی دقت توصیفی است که با بررسی‌های انجام شده توسط نگارندگان، نام عارضه از میان اقلام توصیفی مورد بررسی قرار گرفت. در این مقاله، این معیار هم همانند معیارهای قبلی، از طریق مقایسه دو مجموعه داده مردم‌گستر و استاندارد و تنها بر روی عوارض نظیر انجام خواهد گرفت. یکسان بودن مقادیر توصیفات بین دو مجموعه داده، روشی است که در ابتدا برای ارزیابی دقت نام عارضه استفاده می‌شد. این روش، اسامی راه‌ها را در دو دسته اسامی درست و غلط تقسیم‌بندی می‌کند. در مورد مجموعه داده مردم‌گستر، ممکن است اسامی عوارض املائی نادرستی داشته و یا به صورت اختصاری نوشته شده باشند، درحالی که هر دو مجموعه به یک عارضه یکسان اشاره کنند. در این گونه موارد، اندازه‌گیری بولین مناسب نیست و بهتر است از یک روش جایگزین همانند شباهت متنی<sup>۴</sup> برای این منظور استفاده گردد [۹]. فاصله levenshtein یک الگوریتم شباهت متنی برای تبدیل بین دو رشته به حساب می‌آید [۳۴]. در این روش کمترین مقدار عملگرهایی همانند حذف، اضافه و جابه‌جایی برای تعیین شباهت بین دو رشته محاسبه می‌شوند. در پژوهش حاضر از الگوریتم levenshtein (رابطه ۱) که بیشترین کاربرد را در مباحث تناظر یابی دارد، استفاده شده است.

$$Lev_{a,b}(i,j) = \begin{cases} \max(i,j), & \text{if } \min(i,j) = 0 \\ \min \begin{cases} lev_{a,b}(i-1,j) \\ lev_{a,b}(i,j-1) \\ lev_{a,b}(i-1,j-1) + 1(a_i \neq b_i) \end{cases} & \text{otherwise} \end{cases} \quad (1)$$

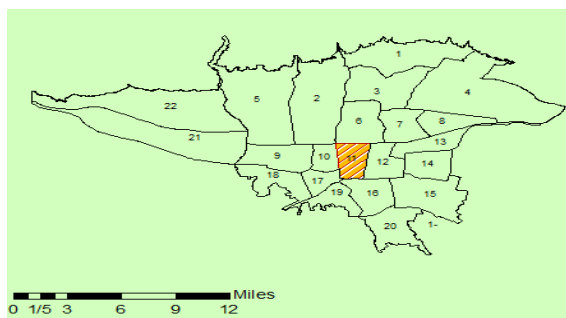
<sup>۵</sup> Overshoot, Undershoot (Dangle Error)  
<sup>۶</sup> Duplicate Error  
<sup>۷</sup> Intersected Roads without a Common Node

<sup>۱</sup> Increasing Buffer  
<sup>۲</sup> Root Mean Square Error  
<sup>۳</sup> Maximum Devotion  
<sup>۴</sup> Text Similarity

برای هر یک از ورودی‌ها سه تابع عضویت مثلثی و خروجی با استفاده از ۵ تابع عضویت تعریف گردید. مرحله سوم؛ بدلیل ناهمگن بودن متغیرهای ورودی ابتدا، هر یک از آن‌ها بین بازه [0-1] نرمال سازی می‌شوند سپس مجموعه قواعد مناسب جهت ساخت مدل استخراج می‌گردند. در این تحقیق، وزن تمام معیارها ثابت در نظر گرفته شده است به همین منظور تمام حالات ممکن برای تعریف قوانین در نظر گرفته می‌شود. مرحله چهارم؛ در این مرحله قوانین تعریف شده روی داده‌ها اعمال شده و برای هر سلول یک خروجی فازی به‌عنوان برازندگی استفاده داده‌های مردم‌گستر به دست خواهد آمد. در این تحقیق از عملگر "و" برای ترکیب متغیرهای هر قانون استفاده شده است. حال برای اینکه خروجی فازی به اعداد معمولی تبدیل شوند، با استفاده از روش مرکز ثقل، ارزش خروجی‌ها غیر فازی می‌گردند.

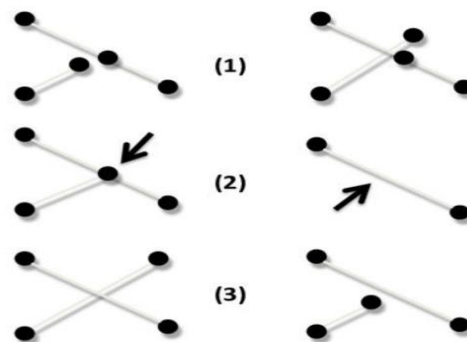
#### ۵- پیاده‌سازی و ارزیابی

در این پژوهش، منطقه ۱۱ شهر تهران به‌عنوان منطقه مطالعاتی انتخاب گردید. این منطقه از مناطقی است که در مرکز شهر تهران واقع شده است. مساحت آن برابر ۱۲٫۶ کیلومترمربع و حدوداً ۸٫۱ درصد از کل مساحت تهران را می‌پوشاند (شکل ۵). داده‌های گردآوری شده در این تحقیق شامل داده‌های شبکه راه در مقیاس ۱:۲۰۰۰، تهیه شده توسط شهرداری تهران به‌عنوان مجموعه داده استاندارد و داده‌های سایت OSM به‌عنوان مجموعه داده مردم‌گستر هستند.



شکل ۵- منطقه ۱۱، منطقه مورد مطالعه

همان‌طور که پیش‌ازین بحث شد، انتخاب عرض بافر مناسب یکی از پیش‌پردازش‌های مهم در این تحقیق به شمار می‌آید. به همین دلیل در یک فرآیند خودکار بافرهایی اطراف داده‌های استاندارد ترسیم می‌شود. با در



شکل ۴- نمایشی از خطاهای توپولوژی: به ترتیب خطای (۱) dangle، (۲) duplicate و (۳) intersect [۳۵]

#### ۴-۴- برآورد برازندگی استفاده

مدل‌های ترکیب متعددی نظیر همپوشانی شاخص<sup>۱</sup>، عملگری بولین<sup>۲</sup>، منطق فازی<sup>۳</sup>، وزن‌دهی شواهد<sup>۴</sup> و الگوریتم ژنتیک<sup>۵</sup> وجود دارند. بدلیل این‌که تحقیق حاضر ابزاری برای مدیریت بحران و بخصوص مسیریابی گروه‌های امدادی تلقی می‌شود. نیازمند روشی منعطف نسبت به اعمال نظرات کارشناسانه در نتیجه حاصل است تا بتواند خروجی مطلوب‌تری بدست آورد. همچنین در تحقیق صورت گرفته توسط Forghani و Delavar [۱۴] عدم قطعیت عوارض مردم‌گستر با استفاده از سیستم خبره فازی تعیین گردید. در این تحقیق، براساس مطالعات پیشین و دلیل یاد شده، استفاده از سیستم خبره فازی پیشنهاد گردید. همچنین از روش حداقل- حداکثر ممدانی، به علت ساختار مؤثر و ساده که بیشترین کاربرد را در مسائل علمی پیدا کرده است، بهره گرفته شد. مراحل طراحی این سیستم به این ترتیب است. مرحله اول؛ در این مرحله ورودی‌ها و خروجی‌های مدل تعیین می‌گردند. به طور مثال معیارهای کاملی شبکه راه، دقت هندسی، دقت توصیفی، کاملی توصیفی و سازگاری منطقی شبکه راه به‌عنوان ورودی‌ها و برازندگی استفاده خروجی این سیستم معرفی شدند. مرحله دوم؛ در این مرحله متغیرهای ورودی و خروجی فازی می‌گردند. در واقع در این مرحله هر یک از ورودی‌ها و خروجی‌ها با تعریف توابع عضویت به متغیرهای زبانی تبدیل می‌شوند. در این تحقیق

۱ Index Overlay  
 ۲ Boolean Operation  
 ۳ Fuzzy logic  
 ۴ Weight of evidence  
 ۵ Genetic Algorithm

جدول ۱- طول عوارض مردم‌گستر و استاندارد و اختلاف طولی عوارض متناظر مردم‌گستر با مجموعه‌های اصلی

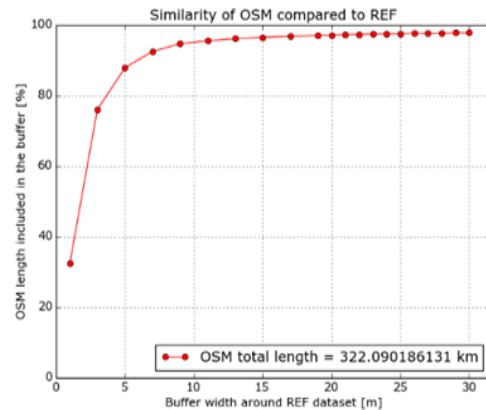
طول عوارض OSM	۳۲۲۰۹۰/۲ متر
طول عوارض Reference	۳۰۹۵۰۹/۳ متر
تفاوت طولی بین OSM و OSM متناظر شده	۱۵۱۷۴/۹ متر
	۴/۷ درصد
تفاوت طولی بین OSM متناظر شده و Reference	۲۵۹۴ متر
	۰/۸ درصد

بعد از اتمام فرآیند تناظر یابی، کل منطقه مورد مطالعه به سلول‌های ۱۰۰ متر در ۱۰۰ متر تقسیم‌بندی شده، و سپس یک همپوشانی روی دو مجموعه داده مردم‌گستر و داده‌های استاندارد متناظر، به منظور حذف سلول‌هایی که هیچ داده‌ای در درون آن وجود ندارد، اعمال می‌گردد. پس از آن، به برآورد هر یک از معیارهای نامبرده می‌پردازیم. برای ارائه دقت برآورد شده هر یک از معیارها، متدهای متفاوتی وجود دارد، به طور مثال می‌توان مقادیر دقت را به صورت کمی و در قالب جدول توصیفی گزارش کرد و یا با استفاده از متغیرهای بصری این میزان را روی نقشه نشان داد [۳۶]. به نظر می‌رسد، ارائه بصری دقت، امکانات اکتشافی<sup>۱</sup> بهتری داشته باشد. بنابراین در این پژوهش، از متغیر بصری "رنگ" برای نمایش مقادیر برآورد شده هر یک از معیارها استفاده شد. به این صورت که ابتدا یک طبقه‌بندی عددی برای هر یک از معیارها صورت گرفت و سپس با توجه به اینکه مقدار دقت به دست آمده برای هر پیکسل در کدام یک از دسته‌ها قرار گرفته، رنگ مخصوص به خود می‌گیرد.

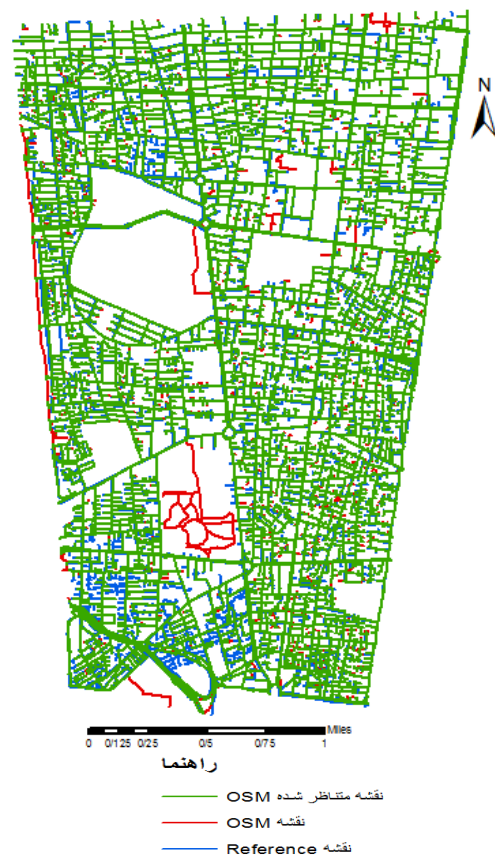
شکل ۸ (الف، ب، ج) برآورد معیارهای کامل بودن، دقت هندسی و سازگاری منطقی به عنوان سنج‌های ارزیابی‌کننده داده‌های هندسی را نشان می‌دهد. کامل بودن با استفاده از رابطه ۲ و دقت هندسی در هر سلول به صورت خودکار محاسبه شده است. درحالی‌که سازگاری منطقی، با استفاده از تعارف قوانین توپولوژی must not overlap و must not dangle می‌آید. از آنجایی که تعداد کل فلگ‌های شناسایی شده برای قانون must not overlap یک بدست آمد، از اعمال این قانون صرف نظر و مجموع دو قانون دیگر در هر پیکسل به عنوان خطاهای توپولوژی برآورد گردید و خروجی آن، به صورت تعداد در هر پیکسل نمایش داده شده است. لازم به ذکر است که هر کدام از معیارها در جدول توصیفی پیکسل مربوطه به صورت جداگانه ذخیره می‌گردد.

<sup>۱</sup> Exploratory

نظر گرفتن سطح اطمینان ۰.۹۷٪، عرض ۱۸ متر، می‌تواند به عنوان عرض بافر مناسب انتخاب گردد (شکل ۶). انتخاب بافر در مراحل تناظر یابی و برآورد دقت هندسی نیز تأثیرگذار است. شکل ۷ داده‌های استاندارد، داده‌های مردم‌گستر و عوارض مردم‌گستری متناظر را نمایش و جدول ۱ طول کل عوارض در هر سه مجموعه داده را نشان می‌دهد.



شکل ۶- عرض بافر اطراف عوارض استاندارد

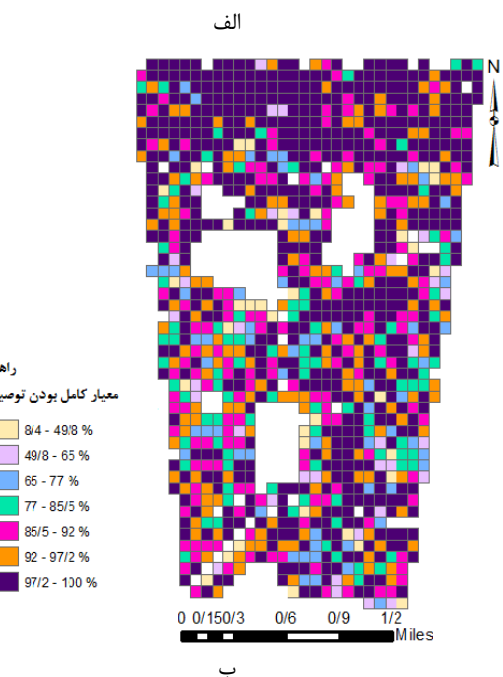
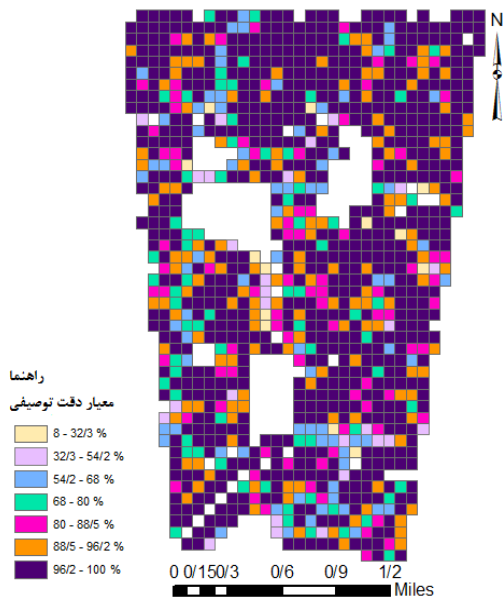


شکل ۷- مجموعه داده مردم‌گستر، استاندارد و مجموعه داده مردم‌گستر متناظر شده

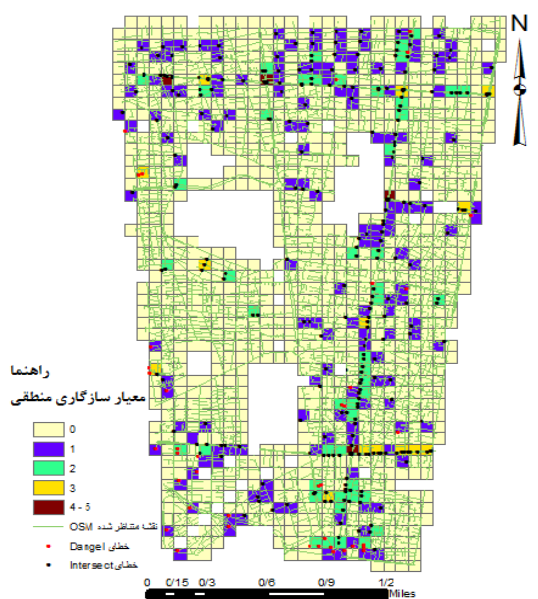
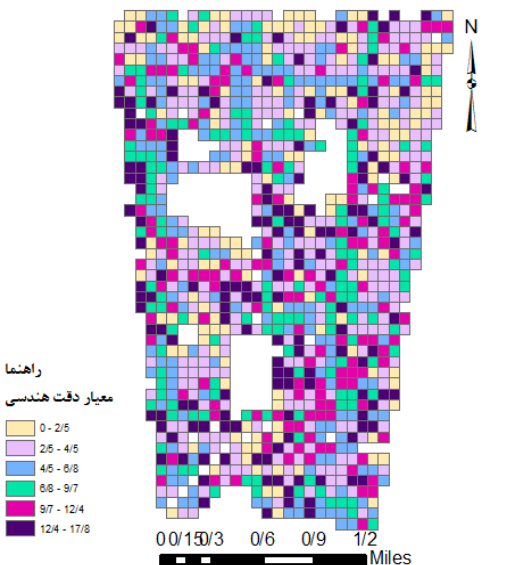
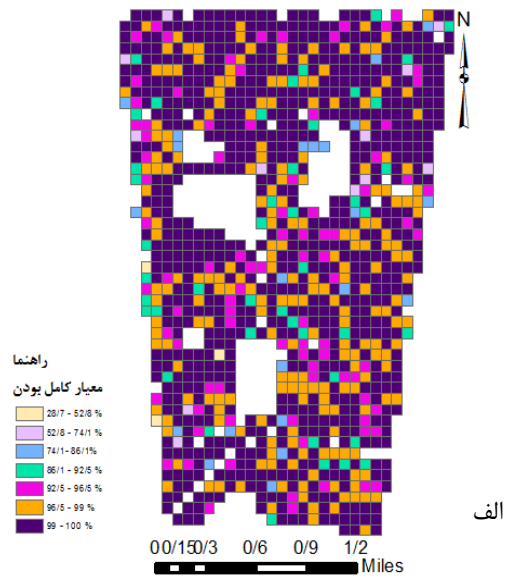
= کامل بودن داده‌های مردم‌گستر در هر پیکسل (درصد)

$$\frac{\text{مجموع طول عوارض REF متناظر}}{\text{مجموع طول عوارض REF در هر سلول}} \times 100 \quad (2)$$

حال نوبت به برآورد معیارهای مربوط به اطلاعات توصیفی می‌رسد. همان‌طور که پیش‌تر اشاره شد، از میان اطلاعات توصیفی، نام عارضه به‌عنوان اطلاعات توصیفی موردبررسی قرار گرفت. رابطه ۴ و ۵ برای برآورد دقت توصیفی و کامل بودن توصیفی استفاده شد [۳۶] و شکل ۹ (الف، ب) برآورد کمی این معیارها را نشان می‌دهد.

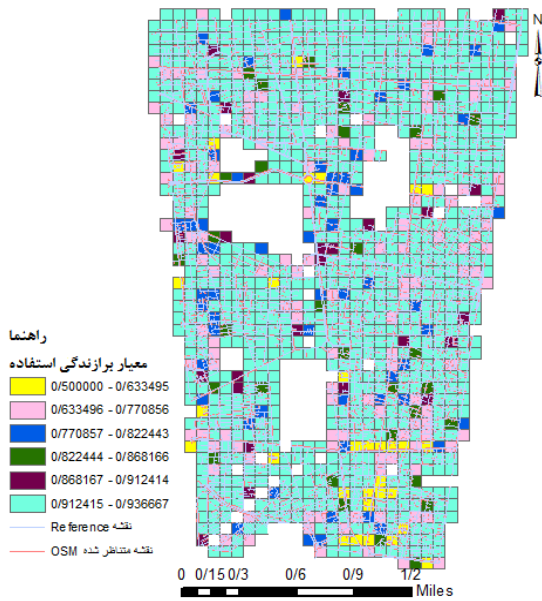


شکل ۹- برآورد معیارهای (الف) دقت توصیفی (ب) کاملی توصیفی نام



شکل ۸- برآورد معیارهای (الف) کامل بودن (ب) دقت هندسی (ج) سازگاری منطقی

نهایتاً با اعمال قوانین، خروجی یا به عبارتی برازندگی استفاده به صورت فازی خواهد بود که به منظور نافازی سازی، از روش سنتروید استفاده گردید و مطابق با شکل ۱۱ خروجی نهایی مقداری بین ۰/۵ تا ۰/۹۳ به دست آمد. همان طور که از شکل ۱۰ و جدول ۳ پیداست، ۷۴/۴ درصد پیکسل‌ها، برازندگی استفاده بین ۰/۹۱ تا ۰/۹۳ دارند. همین طور میانگین برازندگی برای کل منطقه معادل ۰/۸۹ است که یک نتیجه کاملاً ایده آلی است و برای اهدافی نظیر مسیریابی و ناوبری مناسب بنظر می‌رسد.



شکل ۱۱- خروجی برازندگی استفاده

جدول ۳- مقادیر عددی برآورد برازندگی استفاده

طبقه‌بندی مقادیر						برازندگی استفاده
(۰/۹۳ - ۰/۹۱)	(۰/۸۶ - ۰/۸۴)	(۰/۸۲ - ۰/۸۰)	(۰/۷۷ - ۰/۷۴)	(۰/۶۳ - ۰/۶۰)	(۰/۵ - ۰/۴۳)	
۷۴/۴ درصد	۲/۳۹ درصد	۲/۸۵ درصد	۵/۵۲ درصد	۱۱/۹۷ درصد	۲/۸۵ درصد	

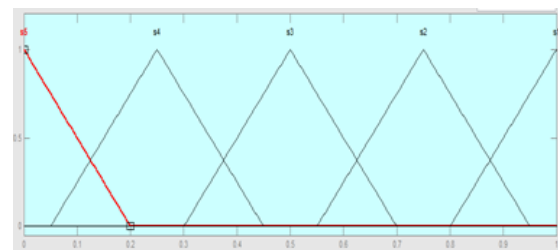
## ۶- ارزیابی

یکی از راهکارهای موجود برای ارزیابی در این تحقیق، تست عملکردی داده‌های OSM در کاربرد مسیریابی است. برای این منظور می‌توان با استفاده از تحلیل شبکه، کوتاه‌ترین مسیر بین مناطق مشخص را با اعمال الگوریتم Dijkstra تعیین و سپس بررسی کرد که آیا کوتاه‌ترین مسیر به دست آمده به لحاظ طولی، در واقعیت موجود نیز کوتاه‌ترین است. بنابراین، آنالیز کوتاه‌ترین مسیر در دو

$$= \frac{\text{مجموع طول عوارض OSM با فاصله ویرایش صفر و یک}}{\text{مجموع طول عوارض OSM دارای نام در هر پیکسل}} \times 100 \quad (۳)$$

$$= \frac{\text{مجموع طول عوارض OSM دارای نام}}{\text{مجموع طول کل عوارض OSM در هر پیکسل}} \times 100 \quad (۴)$$

مرحله آخر، برای برآورد برازندگی استفاده، لازم است ابتدا برای فازی سازی معیارهای ورودی و خروجی توابع عضویت تعریف گردد تا میزان تعلق هر عنصر به مجموعه فازی توسط این توابع تعیین شود. در این پژوهش، از توابع عضویت مثلثی استفاده شد به گونه‌ای که برای هر یک از معیارهای ورودی سه تابع عضویت "مناسب"، "متوسط" و "نامناسب" تعریف و برای خروجی پنج تابع عضویت خیلی خوب (S1)، خوب (S2)، متوسط (S3)، پایین (S4) و خیلی پایین (S5) مطابق شکل ۱۰ تعریف گردید. توابع عضویت بر اساس نظر کارشناسان و اصول ساخت این توابع به دست آمدند. سپس قوانین فازی به صورت کامل و با بررسی تمام حالات ممکن تعریف شد. نمونه‌ای از این قوانین در جدول ۲ نمایش داده شده است.



شکل ۱۰- توابع عضویت خروجی (برازندگی استفاده)

جدول ۲- نمونه‌ای از قوانین فازی تعریف شده

1. If (positional_acc is good) and (completeness is low) and (Attribute_acc is low) and (Att_completeness is low) and (Logical_con is good) then (output1 is s4)
2. If (positional_acc is good) and (completeness is low) and (Attribute_acc is low) and (Att_completeness is low) and (Logical_con is moderate) then (output1 is s4)
3. If (positional_acc is moderate) and (completeness is low) and (Attribute_acc is low) and (Att_completeness is low) and (Logical_con is low) then (output1 is s5)
4. If (positional_acc is good) and (completeness is low) and (Attribute_acc is low) and (Att_completeness is moderate) and (Logical_con is good) then (output1 is s3)
5. If (positional_acc is moderate) and (completeness is moderate) and (Attribute_acc is moderate) and (Att_completeness is good) and (Logical_con is moderate) then (output1 is s3)
6. If (positional_acc is good) and (completeness is low) and (Attribute_acc is good) and (Att_completeness is good) and (Logical_con is good) then (output1 is s1)



با توجه به متوسط برآزندگی ۸۹ درصدی و خروجی 95/72 درصدی از ارزیابی آنالیز کوتاه‌ترین مسیر، می‌توان دریافت که رویکرد پیشنهادی برای ارزیابی برآزندگی داده‌های مردم‌گستر مناسب است و نکته حائز اهمیت تر این است که داده‌های OSM می‌تواند تا حد بسیار زیادی، جوابگوی نیازهای مسئله مسیریابی باشد.

جدول ۴- نتایج ارزیابی

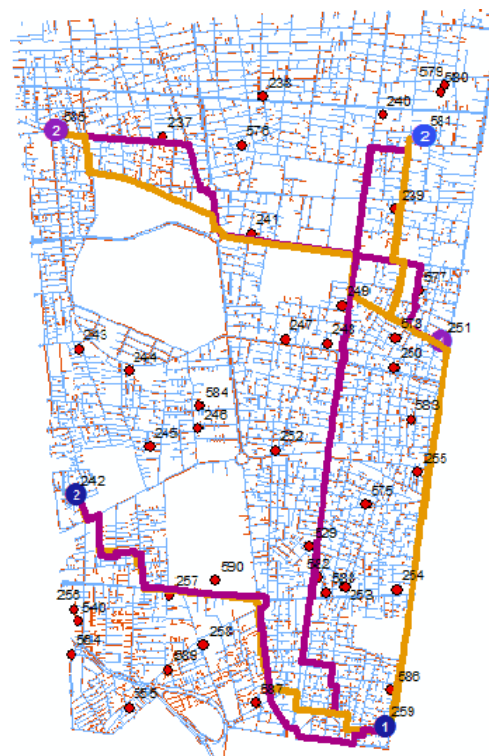
مجموع طول کوتاه‌ترین مسیریاب در OSM	۵۰۵۳۰/۲۹۸۵۸ متر
مجموع طول کوتاه‌ترین مسیریاب در Reference	۴۸۴۵۱/۸۰۷۲۹ متر
اختلاف طولی بین دو مجموعه داده	۲۰۷۸/۴۹۱۲۹ متر + ۴/۲۸ درصد
اختلاف نسبی بین دو مجموعه داده	۹۵/۷۲ درصد

## ۷- نتیجه‌گیری

همان‌طور که پیش‌تر اشاره شد، اطلاعات مکانی مردم‌گستر با وجود داشتن ویژگی‌هایی نظیر به‌روزرسانی، ارزان بودن و در دسترس بودن، گاهی تنها منبع اطلاعاتی در حوزه مدیریت حوادث محسوب می‌شود. اما عدم آگاهی از کیفیت این نوع اطلاعات، موجب گردیده تحقیقات بسیار زیادی برای کسب اطمینان از کیفیت آن‌ها صورت گیرد. با این وجود، مناسب بودن داده‌های مردم‌گستر به‌صورت کلی و در قالب ارائه راهکار بحث شده است. از طرفی مسئله مسیریابی گروه‌های امداد و نجات برای انتقال مجروحین به نقاط امن، یکی از نیازهای اساسی در حوزه مدیریت حوادث است که به عنوان کاربرد موردنظر در این مقاله انتخاب گردید.

برای انجام تحقیق، در ابتدا داده‌های مکانی موردنیاز در این کاربرد شناسایی و سپس معیارهایی برای ارزیابی کیفی آن‌ها با ارجاع به کارهای قبلی استخراج شدند. کامل بودن، دقت هندسی، دقت توصیفی به روش مقایسه با مجموعه داده استاندارد و با استفاده از روش‌های خاص خود برآورد شدند. اما کامل بودن توصیفی و سازگاری منطقی بدون مقایسه با مجموعه استاندارد برآورد شدند. سپس برای شناسایی مناطق با کمترین کیفیت و تقسیم‌بندی منطقه از روش پیکسل مبنای این منظور استفاده گردید، بدین ترتیب هر یک از معیارها در پیکسل‌های ۱۰۰ متر در ۱۰۰ متر برآورد گردیدند.

مجموعه داده مرجع و مردم‌گستر اعمال و سپس اختلاف نسبی در دو مجموعه محاسبه شد. بدین ترتیب، ابتدا ۴۰ نقطه به‌صورت تصادفی در منطقه مورد مطالعه به‌عنوان نقاط مبدأ و مقصد، تولید شده و سپس با تعیین نقطه شروع و پایان یکسان در هر دو مجموعه داده، کوتاه‌ترین مسیر بین آن‌ها تعیین می‌گردد. این فرآیند ۲۰ بار بین ۴۰ نقطه انجام شد. شکل ۱۲ آنالیز کوتاه‌ترین مسیر را بین شش نقطه تصادفی برای دو مجموعه داده نشان می‌دهد. به‌طوری‌که کوتاه‌ترین مسیر در مجموعه مردم‌گستر به رنگ بنفش و در مجموعه استاندارد به رنگ نارنجی نمایش داده شده است. معیار ارزیابی در این تحقیق، طول به‌دست‌آمده از آنالیز کوتاه‌ترین مسیر است، بنابراین اختلاف نسبی مجموع طول کل مسیرهای به‌دست‌آمده در هر دو مجموعه محاسبه می‌گردد. جدول ۴ نتایج حاصل از ارزیابی را نشان می‌دهد. در این جدول اختلاف طولی بین کوتاه‌ترین مسیرها در دو مجموعه محاسبه شده است. طول حاصل از مسیرها در مجموعه داده مردم‌گستر به اندازه ۴/۲۸ درصد بیش از مجموعه استاندارد به دست آمده که بیانگر میزان اختلاف مجموعه داده مردم‌گستر نسبت به استاندارد است و به عبارتی مجموعه داده مردم‌گستر به میزان ۹۵/۷۲ درصد به مجموعه استاندارد شباهت دارد.



شکل ۱۲- تعیین کوتاه‌ترین مسیر در دو مجموعه داده

میانگین، برازندگی استفاده برای کل منطقه ۸۹٪ به دست آمد.

در آخر، برای ارزیابی رویکرد پیشنهادی، از آنالیز کوتاه‌ترین مسیر Dijkstra استفاده شد. بدین ترتیب که کوتاه‌ترین مسیر در دو مجموعه داده تعیین و سپس مجموع طول حاصل موردسنجش قرار گرفت. به طوری که نتایج بیشینه ۴/۲۸٪ داده‌های OSM و شباهت ۹۵٪ نسبت به استاندارد را نشان داد. به طوری که می‌توان نتیجه گرفت که داده‌های OSM تا حد بسیار زیادی برای کاربرد مسیریابی مناسب است و یا به عبارتی جوابگوی نیازهای کاربران در این زمینه را فراهم خواهد کرد.

در مرحله بعد، از یک سیستم خبره فازی به منظور تلفیق معیارهای نامبرده، استفاده شد. در این سیستم با استفاده از روش ممدانی، توابع عضویت برای ورودی و خروجی‌ها و قوانین فازی تشکیل شدند. سپس با استفاده از روش سنتروید، خروجی حاصل فازی زدایی شده و مقداری بین صفر تا یک را به خود اختصاص داد.

در این تحقیق از داده‌های OSM و داده‌های تهیه‌شده توسط شهرداری استفاده شده است. منطقه ۱۱ شهر تهران به‌عنوان محدوده مطالعاتی انتخاب شد. نتایج تحقیق حاکی از آن است که تقریباً ۷۴/۴٪ از کل منطقه مورد مطالعه برازندگی بین ۰/۹۱ تا ۰/۹۳ دارند. و به صورت

## مراجع

- [1] Roche, S., E. Propeck-Zimmermann, and B. Mericskay, GeoWeb and crisis management: Issues and perspectives of volunteered geographic information. *GeoJournal*, 2013. 78(1): p. 21-40.
- [2] Davis, G.B. and M.H. Olson, Management information systems: conceptual foundations, structure, and development. 1984: McGraw-Hill, Inc.
- [3] Goodchild, M.F., in the World of Web 2.0. *International Journal*, 2007. 2: p. 24-32.
- [4] Neis, P., D. Zielstra, and A. Zipf, The street network evolution of crowdsourced maps: OpenStreetMap in Germany 2007–2011. *Future Internet*, 2011. 4(1): p. 1-21.
- [5] Bordogna, G., et al., A linguistic decision making approach to assess the quality of volunteer geographic information for citizen science. *Information Sciences*, 2014. 258: p. 312-327.
- [6] B. Vahedi, A. A. Alesheikh, and S. Honarparvar, Quantitative Assessment of Pragmatic Quality of Volunteered Geographic Information Using Fuzzy Linguistic Quantifiers and OWA Operator. *Journal of Geomatics Science and Technology*, 2014. 3(4): p. 65-76.
- [7] Fan, H., et al., A polygon-based approach for matching OpenStreetMap road networks with regional transit authority data. *International Journal of Geographical Information Science*, 2016. 30(4): p. 748-764.
- [8] Siebritz, L.-A., Assessing the accuracy of openstreetmap data in south africa for the purpose of integrating it with authoritative data. 2014, University of Cape Town.
- [9] Will, J., Development of an automated matching algorithm to assess the quality of the OpenStreetMap road network: a case study in Göteborg, Sweden. Student thesis series INES, 2014.
- [10] Brovelli, M., et al., POSITIONAL ACCURACY ASSESSMENT OF THE OPENSTREETMAP BUILDINGS LAYER THROUGH AUTOMATIC HOMOLOGOUS PAIRS DETECTION: THE METHOD AND A CASE STUDY. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2016. 41.
- [11] Mohammadi, N. and M. Malek, Artificial intelligence-based solution to estimate the spatial accuracy of volunteered geographic data. *Journal of Spatial Science*, 2015. 60(1): p. 119-135.
- [12] Mohammadi, N. and M. Malek, VGI and Reference Data Correspondence Based on Location-Orientation Rotary Descriptor and Segment Matching. *Transactions in GIS*, 2015. 19(4): p. 619-639.
- [13] Forghani, M. and M.R. Delavar, A quality study of the OpenStreetMap dataset for Tehran. *ISPRS International Journal of Geo-Information*, 2014. 3(2): p. 750-763.
- [14] Girres, J.F. and G. Touya, Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS*, 2010. 14(4): p. 435-459.
- [15] Helbich, M., et al., Comparative spatial analysis of positional accuracy of OpenStreetMap and proprietary geodata. *Proceedings of GI\_Forum*, 2012: p. 24-33.
- [16] Ciepluch, B., P. Mooney, and A.C. Winstanley, Building generic quality indicators for OpenStreetMap. 2011.

- [17] Ali Abbaspour, R. and S.H. Hossaini, A Validation Approach for OSM Roads Information without Using Authorized Information Based on the Other OSM Information. *Journal of Geomatics Science and Technology*, 2016. 6(2): p. 1-11.
- [18] Vandecasteele, A. and R. Devillers ,Improving volunteered geographic information quality using a tag recommender system: the case of OpenStreetMap, in *OpenStreetMap in GIScience*. 2015, Springer. p. 59-80.
- [19] Eshghi, M. and A.A. Alesheikh, Introducing a Conceptual Model to Improve the Quality of Storage of Volunteered Geographic Information: In the Field of "Fitness-for-Use" Indicator. *Journal of Geomatics Science and Technology*, 2016. 6(1): p. 19-32.
- [20] Mondzech, J. and M. Sester, Quality analysis of OpenStreetMap data based on application needs. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 2011. 46(2): p. 115-125.
- [21] Poser, K. and D. Dransch, Volunteered geographic information for disaster management with application to rapid flood damage estimation. *Geomatica*, 2010. 64(1): p. 89-98.
- [22] Hung, K.-C., M. Kalantari, and A. Rajabifard, Methods for assessing the credibility of volunteered geographic information in flood response: A case study in Brisbane, Australia. *Applied Geography*, 2016. 68 :p. 37-47.
- [23] Devillers, R. and R. Jeansoulin, Spatial data quality: concepts. *Fundamentals of spatial data quality*, 2006: p. 31-42.
- [24] ISO, I., 19113—Geographic Information—Quality Principles. International Organization for Standardization, 2002.
- [25] Zadeh, L.A., Fuzzy sets. *Information and control*, 1965. 8(3): p. 338-353.
- [26] Vahidnia, M.H., A.A. Alesheikh, and A. Alimohammadi, Hospital site selection using fuzzy AHP and its derivatives. *Journal of environmental management*, 2009. 90(10): p. 3048-3056.
- [27] Koukoletsos, T., A Framework for Quality Evaluation of VGI linear datasets. 2012, UCL (University College London).
- [28] Graser, A., M. Straub, and M. Dragaschnig, Towards an open source analysis toolbox for street network comparison: Indicators, tools and results of a comparison of OSM and the official Austrian reference graph. *Transactions in GIS*, 2014. 18(4): p. 510-526.
- [29] Barron, C., P. Neis, and A. Zipf, A comprehensive framework for intrinsic OpenStreetMap quality analysis. *Transactions in GIS* ,2014 .18(6)p. 877-895.
- [30] Brovelli, M.A., et al., Towards an automated comparison of OpenStreetMap with authoritative road datasets. *Transactions in GIS*, 2017. 21(2): p. 191-206.
- [31] Zielstra, D. and A. Zipf. A comparative study of proprietary geodata and volunteered geographic information for Germany. in *13th AGILE international conference on geographic information science*. 2010.
- [32] Mustière, S. and T. Devogele, Matching networks with different levels of detail. *GeoInformatica*, 2008. 12(4): p. 435-453.
- [33] Goodchild, M.F. and G.J. Hunter, A simple positional accuracy measure for linear features. *International Journal of Geographical Information Science*, 1997. 11(3): p. 299-306.
- [34] Navarro, G., A guided tour to approximate string matching. *ACM computing surveys (CSUR)*, 2001. 33(1): p. 31-88.
- [35] Neis, P., D. Zielstra, and A. Zipf, Comparison of volunteered geographic information data contributions and community development for selected world regions. *Future Internet*, 2013. 5(2): p. 282